

SETEMBRO DE 2020

RELATÓRIO TÉCNICO SOBRE FERRAMENTAS PARA A ELABORAÇÃO DE PLANOS DE GESTÃO DE DADOS

VERSÃO

Título: Relatório Técnico sobre Ferramentas para a Elaboração de Planos de Gestão de Dados

Autores: Pedro Príncipe, Paula Moura, André Vieira, Filipa Pereira

Contribuição: Eloy Rodrigues, João Mendes Moreira, Ana Portugal Melo, João Nuno Ferreira e José Carvalho

Versão: 1

Distribuição: [Equipa de Projeto e disponibilizado publicamente]

Data de Criação: 13 de maio de 2020

Última Atualização: 5 de outubro de 2020

CONTROLO DE REVISÕES AO DOCUMENTO

Revisão	Data	Alterado por	Alterações
1	22 de setembro	Filipa Pereira	Renomeação do capítulo 2.1 das aplicações
2	22 de setembro	André Vieira	Atualização de logos na capa e índice geral

ÍNDICE

ENQUADRAMENTO	4
PLANOS DE GESTÃO DE DADOS	5
O que é e para que serve um Plano de Gestão de Dados	5
Processo de elaboração de Planos de Gestão de Dados	7
Planos de Gestão de Dados acionáveis	8
FERRAMENTAS PARA A ELABORAÇÃO DE PLANOS DE GESTÃO DE DADOS	10
Identificação e descrição das aplicações	10
DMPonline	10
ARGOS - OpenAIRE	12
OpenDMP (software do ARGOS)	14
Data Stewardship Wizard	15
RDMO - Research Data Management Organiser	17
DMPtool	19
ezDMP - NSF	20
easyDMP	22
ARIADNE	23
Análise comparativa de funcionalidades, requisitos e custos	24
Análise de funcionalidades e características	25
Identificação e validação de requisitos base	27
Informação sobre custos	29
ESTUDOS, INICIATIVAS E PROJETOS EM PORTUGAL SOBRE PGDS	31
SERVIÇOS NACIONAIS PARA PGDS	33
PaGoDa - Espanha	33
dmponline.be - Bélgica	34
DMP OPIDoR - França	35
DMP Norway - Noruega (NSD)	35
DMP Online - Dinamarca (DeiC)	36
RDMO - Alemanha	37
CONCLUSÕES	38

ENQUADRAMENTO

No âmbito do plano de atividades na área de gestão de dados de investigação, a Fundação para a Ciência e Tecnologia (FCT), através da Unidade de Computação Científica Nacional (FCCN), pretende disponibilizar à comunidade científica e de ensino superior um sistema de planos de gestão de dados de investigação.

Pretende-se assegurar que os investigadores elaborem e atualizem Planos de Gestão de Dados (PGD) relacionados com os seus projetos, de acordo com a “Política sobre a gestão e partilha de dados e outros produtos resultantes de investigações financiadas pela FCT”¹.

No atual contexto, os Princípios orientadores para a implementação de uma Política Nacional de Ciência Aberta² definem o princípio do “acesso aberto aos dados científicos resultantes de investigação financiada por fundos públicos”, bem como a “garantia da preservação das publicações e dados científicos, por forma a permitir a sua reutilização e o acesso continuado”.

Por outro lado, as Recomendações da Comissão Europeia, de 25 de abril de 2018, sobre o acesso à informação científica e a sua preservação³, referem que os Estados-Membros deverão definir e aplicar políticas claras para a gestão dos dados resultantes de investigação financiada por fundos públicos. Consideram também importante que o “planeamento da gestão de dados se torne uma prática científica normalizada numa fase precoce do processo de investigação quando os dados são gerados ou recolhidos, nomeadamente através da exigência de planos de gestão de dados”.

O presente relatório técnico visa avaliar as ferramentas de gestão de planos de dados de investigação atualmente disponíveis a nível global, de forma a selecionar a que melhor se enquadra no serviço nacional que virá a ser prestado. Esta avaliação integrará vários aspetos considerados relevantes, designadamente:

- A abertura do software, isto é, se é *Open Source* ou proprietário;
- A extensão da comunidade de utilizadores;
- A disponibilidade de modelos para financiadores;
- Os casos de uso;
- A disponibilização de suporte e apoio técnico, bem como formação;
- Os respetivos planos de custos.

¹ Política sobre a gestão e partilha de dados e outros produtos resultantes de investigações financiadas pela FCT. Disponível em;

https://www.fct.pt/documentos/PoliticaAcessoAberto_Dados.pdf

² Resolução do Conselho de Ministros n.º 21/2016. Disponível em:

<https://data.dre.pt/eli/resolconsmin/21/2016/04/11/p/dre/pt/html>

³ Recomendação (UE) 2018/790 da Comissão, de 25 de Abril de 2018, sobre o acesso à informação científica e a sua preservação. Disponível em:

<https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32018H0790&from=EN>

Esta informação permitirá preparar um resumo comparativo das várias funcionalidades disponíveis, facilitando a análise do cumprimento dos requisitos considerados relevantes para o serviço a prestar à comunidade pela FCT.

Serão também abordados os estudos e iniciativas de âmbito nacional sobre PGDs.

Por último, consideramos igualmente importante avaliar os serviços de PGDs que já estão implementados em alguns países de referência, no domínio da gestão de dados de investigação, tais como: Espanha, Bélgica, França, Noruega, Dinamarca e Alemanha.

PLANOS DE GESTÃO DE DADOS

O que é e para que serve um Plano de Gestão de Dados

A gestão dos dados de investigação é fundamental para garantir a qualidade dos dados em todo o ciclo de vida da investigação, viabilizando sempre que possível a sua reutilização. O PGD é um elemento chave para uma boa gestão no decorrer de um projeto de investigação. Descreve todo o ciclo de vida dos dados, desde a sua recolha, processamento e tratamento, até à disponibilização, partilha e acesso.

Um PGD visa especificar todas as questões ligadas com a produção, arquivo, preservação, descrição, partilha e disseminação dos dados gerados no âmbito de um projeto de investigação. Através de um plano de gestão de dados é possível compreender com clareza quais e como é que os dados vão ser recolhidos, tratados e disponibilizados durante e após a finalização de um projeto, bem como a identificação dos intervenientes no processo, a definição do nível de acesso aos dados e a forma como poderão ser reutilizados noutros contextos.

Os Planos de Gestão de Dados são um contributo essencial para um projeto de investigação de qualidade:

- Permitem encontrar e compreender os dados quando precisar de usá-los;
- Garantem continuidade após a saída de elementos da equipa do projeto, ou na entrada de novos investigadores;
- Evitam duplicação de trabalho – ex.: repetir recolha de dados;
- Dados de apoio a publicações são mantidos, permitindo a validação de resultados;
- Promovem a partilha de dados, o que leva a maior colaboração e avanços na investigação;
- A investigação torna-se mais visível e ganha maior impacto;
- Outros investigadores podem citar os seus dados.

Os PGDs são frequentemente requeridos como parte dos acordos de subvenção dos financiadores de ciência, mas são úteis em todos os projetos sempre que os investigadores estão a gerar dados. Torna-se assim necessário considerar as especificações que os financiadores de ciência e inovação estabelecem para os seus beneficiários, sendo cada vez mais comum a exigência de apresentação de um PGD na fase da proposta do projeto.

As principais razões para a criação de um PGD apresentadas pelos financiadores são:

- Promover boas práticas de gestão de dados que assegurem a devida documentação e preservação dos conjuntos de dados, evitando perdas e custos associados; e
- Melhorar e maximizar o acesso e a reutilização dos dados de investigação, garantindo que os resultados de projetos com financiamento público sejam disponibilizados para reutilização e verificação pela sociedade em geral.

Um PGD é um documento que define, relativamente aos dados:

- Como serão gerados;
- Como serão documentados;
- Quem os poderá aceder;
- Onde e como serão guardados;
- Quem fica responsável pela sua preservação;
- Se e como podem ser partilhados e preservados.

Um PGD apoia no planeamento e organização da recolha de dados, fazendo com que se pense com cuidado nas questões que surgirão quando se começar a produzir e recolher dados. Um PGD documenta as principais atividades do ciclo de vida dos dados de investigação, como recolha, descrição, preservação e acesso ou pesquisa de dados. Essa documentação é crucial para a reprodutibilidade dos resultados da investigação, que é um princípio fundamental da investigação científica. Um PGD fornece detalhes valiosos relativos ao processo como os dados serão preservados a longo prazo, como e onde o investigador disponibilizará os dados para partilha e reutilização, e quando esse acesso aberto será permitido.

A elaboração de um PGD contribui para prevenir ou reduzir a probabilidade de contratempos, como perda de dados, erros de dados e usos desadequados dos mesmos. Com efeito, um PGD promove uma melhor documentação e comunicação dos dados e ainda uma maior responsabilização pelos mesmos.

Apresentar um PGD reflete o entendimento de que os dados produzidos e recolhidos têm valor intrínseco, sendo por isso mesmo fonte de atribuição, passíveis de crédito aos seus autores, devendo ser referenciados e devidamente citados por outros que os reutilizem.

O PGD não é um documento fixo, evolui e ganha maior precisão e substância à medida que o projeto se desenvolve, uma vez que nem todos os dados ou usos potenciais são claros desde o princípio.

Processo de elaboração de Planos de Gestão de Dados

A elaboração de PGDs pode ser agilizada com o recurso a ferramentas já existentes, disponibilizadas por diferentes entidades. Para facilitar o trabalho dos investigadores, algumas destas ferramentas já providenciam os planos de acordo com os requisitos dos financiadores. Todas são passíveis de serem utilizadas, parametrizadas e ajustadas à medida do projeto que os investigadores se encontrem a gerir.

Tópicos mais comuns nos PGDs e questões que os investigadores devem procurar responder no processo de elaboração de um PGD (identificação baseada nos elementos definidos pela Science Europe)⁴:

- 1) Descrição dos dados e recolha ou reutilização dos dados existentes:
 - a) Como novos dados serão recolhidos ou produzidos e/ou como os dados existentes serão reutilizados?
 - b) Que dados (por exemplo, tipos, formatos e volumes) serão recolhidos ou produzidos?
- 2) Documentação e qualidade dos dados:
 - a) Que metadados e documentação (por exemplo, a metodologia de recolha de dados e o modo de organização dados) acompanhará os dados?
 - b) Que medidas de controlo da qualidade dos dados serão utilizadas?
- 3) Armazenamento e backup durante o processo de investigação:
 - a) Como os dados e metadados serão armazenados durante o processo de investigação?
 - b) Como a segurança dos dados e a proteção de dados sensíveis serão tratadas durante a investigação?
- 4) Requisitos legais e éticos, códigos de conduta:

⁴ Science Europe. Practical Guide to the International Alignment of Research Data Management. Disponível em:
https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf

- a) Se forem tratados dados pessoais, como será assegurado o cumprimento da legislação relativa aos dados pessoais e à proteção dos dados?
 - b) Como serão geridas outras questões jurídicas, como os direitos de propriedade intelectual? Que legislação é aplicável?
 - c) Como serão tidas em conta as possíveis questões éticas, e códigos de conduta seguidos?
- 5) Partilha de dados e preservação a longo prazo:
- a) Como e quando os dados serão disponibilizados em acesso aberto? Há possibilidade de restrições à partilha de dados ou motivos de embargo?
 - b) Como serão selecionados os conjuntos de dados para preservação e onde será assegurada a preservação a longo prazo (por exemplo, repositórios de dados ou arquivos de dados)?
 - c) Que métodos ou ferramentas de software serão necessários para aceder e usar os dados?
 - d) Como é que será assegurado o registo de identificador único e persistente (como o DOI) para cada conjunto de dados?
- 6) Responsabilidades e recursos em matéria de gestão de dados:
- a) Quem (por exemplo, função, cargo e instituição) será responsável pela gestão dos dados (ou seja, o administrador de dados)?
 - b) Que recursos (por exemplo, financeiros e de tempo) serão dedicados à gestão de dados e à garantia de que os dados são compatíveis com os princípios FAIR (**F**indable, **A**ccessible, **I**nteroperable, **R**e-usable)?

Planos de Gestão de Dados acionáveis

Os PGDs são documentos de texto, normalmente disponibilizados em formato PDF, Word (DOCX ou DOC) ou RTF, que acompanham propostas ou resultados de projetos de investigação. Estes documentos descrevem os dados que serão gerados, as ferramentas a usar e as políticas a seguir, sendo muitas vezes vistos como meros documentos administrativos, criados para cumprir requisitos de entidades financiadoras. No entanto, os planos podem e devem ser instrumentos de apoio para a gestão dos dados, mas também para aferir as responsabilidades de curadoria ao longo do ciclo de vida dos dados, portanto, parte integrante do processo de investigação.

Neste contexto, a integração dos PGDs nos fluxos de trabalho dos diferentes intervenientes e componentes do ecossistema de dados de investigação é reconhecidamente um dos principais desafios para o desenvolvimento de planos para a gestão dos dados em projetos de investigação. Melhorar a interação entre os componentes envolvidos no ciclo de vida dos

dados e facilitar a troca de informações entre ferramentas e sistemas, é possível a partir da incorporação dos próprios PGDs nos fluxos de trabalho existentes. Para esse efeito, será importante desenvolver serviços que permitam que partes do PGD sejam geradas ou incorporadas automaticamente para melhorar a qualidade da informação e reduzir esforço administrativo, bem como, promover uma maior interoperabilidade entre sistemas, como por exemplo, a identificação dos repositórios onde se prevê o depósito dos dados produzidos.

Os PGDs podem ser documentos vivos, de maior interação com máquinas e mais programáveis. É neste contexto que surge a referência a planos acionáveis - “Actionable Data Management Plans”. O conceito de planos acionáveis - Actionable Data Management Plans - surgiu recentemente, no quadro da comunidade da IDCC⁵, sendo depois estudado na RDA, onde foram definidas as 10 regras para *machine-actionable DMPs*⁶:

- 1) Integrar os PGDs com os fluxos de trabalho de todos os intervenientes no ecossistema de dados de investigação;
- 2) Permitir que os sistemas automatizados atuem em representação dos intervenientes no ecossistema de dados de investigação;
- 3) Elaborar políticas que integrem os sistemas e não apenas as pessoas;
- 4) Descrever, tanto para pessoas como para sistemas, os componentes do ecossistema de gestão de dados;
- 5) Utilizar identificadores persistentes e vocabulários controlados;
- 6) Seguir um modelo de dados comum para planos de gestão de dados acionáveis;
- 7) Disponibilizar PGDs para utilização por parte de pessoas e sistemas;
- 8) Apoiar a avaliação e monitorização da gestão de dados;
- 9) Tornar os PGDs documentos atualizáveis, vivos e com possibilidade de *versionamento*;
- 10) Tornar os PGDs publicamente disponíveis.

Entretanto, no contexto da RDA, foi já apresentado o modelo das normas a seguir para os PGDs acionáveis - *DMP Common Standard for machine-actionable Data Management Plans*⁷. As ferramentas disponíveis online para apoiar a elaboração de planos devem, por isso mesmo, alinhar com as especificações propostas pelo grupo de trabalho da RDA dedicado aos “DMP Common Standards”⁸.

⁵ International Digital Curation Conference

⁶ Miksa, Tomasz, Simms, Stephanie, Mietchen, Daniel, & Jones, Sarah. (2018). *Ten simple rules for machine-actionable data management plans* (Version preprint). <http://doi.org/10.5281/zenodo.1434938>

⁷ RDA, DMP Common Standards WG. *Machine-actionable DMPs - take them and use them!*. Disponível em: <https://www.rd-alliance.org/system/files/documents/2019-RDA-DMP-Helsinki.pdf>

⁸ DMP Common Standards WG: <https://www.rd-alliance.org/groups/dmp-common-standards-wg>

FERRAMENTAS PARA A ELABORAÇÃO DE PLANOS DE GESTÃO DE DADOS

Nesta secção do relatório apresentam-se as principais ferramentas ou aplicações online para a elaboração de planos de gestão de dados, disponibilizadas no âmbito da ação de organismos nacionais de apoio à gestão de dados, ou financiadores de ciência e inovação, ou no contexto de iniciativas de infraestruturas europeias de investigação.

Identificação e descrição das aplicações

DMPonline

	<p>Designação DMPonline https://dmponline.dcc.ac.uk/</p> <p>Proprietário Digital Curation Centre (DCC)</p>
<p>Breve descrição do serviço</p> <p>O DMPonline é uma ferramenta <i>online</i> que apoia os investigadores na preparação de planos de gestão de dados. Esta ferramenta integra os modelos mais recentes de financiadores e também orientações das melhores práticas para a criação de planos de gestão de dados de qualidade.</p> <p>O DMPonline apresenta três funções principais:</p> <ul style="list-style-type: none"> • Ajudar a criar e manter diferentes versões dos planos de gestão de dados (versão de uso na candidatura, versão a ser desenvolvida durante o projeto e versão final); • Fornecer orientações úteis em questões de gestão de dados e como atender aos requisitos de investigação dos financiadores; • Exportar planos úteis numa variedade de formatos. <p>É importante referir que o Digital Curation Centre (DCC) está a colaborar com a EUDAT, através da parametrização do DMPonline para integração dos serviços existentes B2SHARE (https://b2share.eudat.eu/) e B2DROP (https://b2drop.eudat.eu/), assegurando a partilha e o depósito seguro de dados.</p>	

Informação do Software

Software *Open Source* baseado no ROADmap, gerido pelo DCC em parceria com a California Digital Library.

Comunidade de utilizadores

A comunidade que utiliza o DMPonline é bastante alargada:

- Cerca de 17.600 utilizadores;
- 203 organizações;
- Presente em 89 países;
- Total de 23.000 planos de gestão de dados (informação recolhida a 19/05/2020).

A ferramenta está igualmente integrada no GitHub:

https://github.com/DigitalCurationCentre/DMPonline_v4

Modelos para Financiadores

Esta ferramenta disponibiliza modelos de PGDs para financiadores, permitindo a criação de novos modelos de acordo com os requisitos da organização ou financiador.

Casos de uso

O DMPonline dá apoio e suporte a várias instituições universitárias britânicas através do *Higher Education Funding Council of England*.

O apoio institucional através do Research Data Service pode ser consultado em:

<https://www.ed.ac.uk/information-services/research-support/research-data-service/before/writing-a-data-management-plan>.

Informação de suporte

Do ponto de vista de apoio, existem diversas instâncias disponíveis, destacando-se:

- Guia de apoio para a preparação de planos de gestão de dados: <https://www.dcc.ac.uk/basic-page/dmps>
- Apoio na utilização da aplicação:
 - <https://dmponline.dcc.ac.uk/help>
- Apoio a administradores do software:
 - <https://github.com/DigitalCurationCentre/roadmap/wiki/Help-for-administrators>

Planos de Custos

Informação não disponível publicamente. Existe sim alguma informação referente aos modelos de subscrição, que poderá ser consultada em:

<https://www.dcc.ac.uk/blog/dmponline-subscription-model-revisited>

ARGOS - OpenAIRE

**Designação**ARGOS | <https://argos.openaire.eu/>**Proprietário**

OpenAIRE

Breve descrição do serviço

O ARGOS é um serviço aberto e extensível que simplifica a gestão, validação, monitorização e manutenção de PGDs. Permite aos investigadores, gestores, supervisores, ou outros intervenientes, a criação de PGDs que podem ser livremente interoperáveis entre infraestruturas, para ir ao encontro dos aspetos específicos do processo de gestão de dados de acordo com as intenções e compromissos dos proprietários dos dados.

É uma ferramenta online de apoio a processos automatizados de criação, gestão, partilha e ligação do PGD com os resultados de investigação correspondentes. Este serviço tem como objetivo a disponibilização de uma plataforma aberta para o Planeamento de Gestão de Dados que aborde as melhores práticas FAIR e de Abertura. Fá-lo através da aplicação de normas para PGDs acionáveis por máquina, como definido pela RDA e interagindo com investigadores, comunidades de investigação e financiadores, para ir ao encontro das suas necessidades.

Disponibiliza um ambiente flexível e um interface de fácil utilização para os utilizadores navegarem e utilizarem.

Principais características:

- Diferencia as descrições de PGDs e Datasets, podendo os PGDs ter mais do que uma descrição de Dataset;
- Um PGD pode incluir descrições de Datasets de mais do que um modelo de financiador;
- Permite a redação colaborativa;
- Suporta a exportação no formato JSON;
- Disponibiliza DOIs e versionamento de PGD (via Zenodo);
- Financiadores têm a possibilidade de adicionar, atualizar e modificar modelos de PGDs.

Informação do Software

Open Source.

Código e documentação:

<https://issue.openaire.research-infrastructures.eu/projects/openaire/wiki/Argos>

Comunidade de utilizadores

Roadmap público: <https://trello.com/b/x49lylnK/argos>

Modelos para Financiadores

Este serviço disponibiliza modelos de PGDs para financiadores e permite a criação de novos à medida dos requisitos da organização ou financiador.

Casos de uso

132 Utilizadores, 85 PGDs e 91 Descrições de Datasets (informação recolhida a 26/05/2020). O ARGOS foi apresentado pela Comissão Europeia como um dos serviços disponíveis para suporte à gestão de dados de investigação em torno do COVID-19⁹.

A FECYT (Fundação Espanhola para a Ciência e a Tecnologia) utiliza o ARGOS.

O serviço encontra-se também em disseminação por vários países Europeus, pela rede de secretariados nacionais do OpenAIRE, através de sessões de formação, estando ainda a contribuir para o seu desenvolvimento através do envio de feedback com base nas experiências de utilização. Uma das contribuições desta rede é também a tradução da interface do serviço para outros idiomas, como o espanhol e o grego, estando ainda em processo de tradução para húngaro, alemão e português.

Informação de suporte

Guia de utilização do serviço: <https://www.openaire.eu/argos-guide>.

Informação no catálogo de serviços do OpenAIRE:


<http://catalogue.openaire.eu/service/openaire.argos>

Planos de Custos

Informação não disponível publicamente.

⁹ Guidelines for open access to publications, data and other research outputs. Disponível em: https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-guidelines-oa-covid-19_en.pdf

OpenDMP (software do ARGOS)

	<p>Designação OpenDMP https://gitlab.eudat.eu/dmp/OpenAIRE-EUDAT-DMP-service-pilot</p> <p>Proprietário OpenAIRE, EUDAT</p>
<p>Breve descrição do serviço</p> <p>O OpenDMP é o software do serviço ARGOS, desenvolvido num esforço conjunto do OpenAIRE e EUDAT, disponibilizando as mesmas funcionalidades apresentadas pelo ARGOS.</p> <p>O OpenDMP permite a instalação de uma instância local na instituição que pretenda implementar o serviço.</p> <p>Características:</p> <ul style="list-style-type: none"> • <i>PGD Machine-actionable</i>, alinhado com normas globais: <i>RDA DMP Common Standard for machine-actionable Data Management Plans</i> (https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard). <p>Tecnologia:</p> <ul style="list-style-type: none"> • Frontend: <ul style="list-style-type: none"> - Single Page Application (SPA); - Angular / Typescript; - Material Design / Responsive. • Backend: <ul style="list-style-type: none"> - REST backend services; - Java/Spring Stack; - PostgreSQL RDBMS. 	

Informação do Software

Open Source: Apache 2.

Comunidade de utilizadores

Roadmap público: <https://trello.com/b/x49lylnK/argos>

GitLab: <https://gitlab.eudat.eu/dmp/OpenAIRE-EUDAT-DMP-service-pilot>

Código e documentação:

<https://issue.openaire.research-infrastructures.eu/projects/openaire/wiki/Argos>

Modelos para Financiadores

Esta ferramenta disponibiliza modelos de PGDs para financiadores e permite a criação de novos à medida dos requisitos da organização ou financiador.

Casos de uso

Projetos que implementaram o OpenDMP: easyDMP (EUDAT) e ARGOS (OpenAIRE).

Estão também a decorrer negociações para a implementação do serviço em algumas organizações a nível nacional, como por exemplo na Hungria.

Informação de suporte

Código e documentação:

<https://issue.openaire.research-infrastructures.eu/projects/openaire/wiki/Argos>

Planos de Custos

Custos ainda não definidos, no entanto, serão definidos de acordo com o nível de suporte requerido pela organização. Quanto à preparação de modelos *machine-actionable*, ainda a trabalhar com um custo fixo, sendo atualmente gratuito.

Data Stewardship Wizard

	<p>Designação Data Stewardship Wizard https://ds-wizard.org</p> <p>Proprietários ELIXIR CZ e ELIXIR NL</p>
<p>Breve descrição do serviço</p> <p>O Data Stewardship Wizard é uma ferramenta online que resulta de uma parceria entre o ELIXIR CZ e o ELIXIR NL. Esta aplicação visa dar apoio aos investigadores na criação e manutenção de planos de gestão de dados e no cumprimento dos princípios FAIR. Esta ferramenta permite igualmente a disponibilização de apoio por parte de especialistas do ELIXIR.</p> <p>O Data Stewardship Wizard integra a iniciativa #activeDMPs e apresenta a vertente <i>machine-actionable</i>.</p>	

Informação do Software

Open Source.

Comunidade de utilizadores

As instituições que utilizam o Data Stewardship Wizard estão maioritariamente relacionadas com o domínio das Ciências da Vida. Alguns exemplos de utilizadores incluem:

- SciLifeLab;
- INESC-ID Lisboa;
- BioData.pt;
- Fundação Calouste Gulbenkian;
- Universidade de Lisboa;
- University of Tromsø.

A ferramenta está integrada no GitHub: <https://github.com/ds-wizard>.

Modelos para Financiadores

Esta ferramenta disponibiliza modelos para financiadores, permitindo a criação de novos modelos de acordo com os requisitos da organização ou financiador.

Casos de uso

Os casos de uso desta ferramenta são feitos com base nas experiências das instalações já em curso, designadamente no que respeita à área das Ciências da Vida.

Informação de suporte


Do ponto de vista de suporte, as orientações são providenciadas através do GitHub (<https://github.com/ds-wizard>) e incluem:

- Orientações sobre os planos de gestão de dados;
- Criação e customização de modelos.

Planos de Custos

Informação não disponível publicamente.

RDMO - Research Data Management Organiser

	<p>Designação RDMO - Research Data Management Organiser https://rdmorganiser.github.io/en/</p> <p>Proprietário A RDMO é financiada pela Fundação Alemã de Investigação (DFG) - Deutsche Forschungsgemeinschaft</p>
<p>Breve descrição do serviço</p> <p>É uma aplicação web, em código aberto, disponível no GitHub, estando ainda em fase de desenvolvimento (2.^a fase do projeto).</p> <p>A Research Data Management Organiser (RDMO) visa prestar apoio no planeamento sistemático, na organização e implementação da gestão de dados de investigação ao longo de um projeto. Complementarmente dá a possibilidade de extração de um plano de gestão de dados, sob a forma de um documento de texto, de acordo com os requisitos do financiador.</p>	

No [arranque do projeto¹⁰](#), para o desenvolvimento da arquitetura e do modelo de base de dados, recorreram à experiência existente e ao trabalho já desenvolvido, fazendo uso do código *Open Source* do DMPonline e do DMPTool.

Entre as principais funcionalidades nesta 1.^a fase do projeto estava incluído:

- Instância central para todas as informações relacionadas com a gestão de dados de investigação de um projeto;
- Modelos adaptáveis para questionários e relatórios;
- Personalização fácil do aspeto, versionamento da informação;
- Suporte multilingue integrado;
- Modelos para vários fins, por exemplo, PGD para financiador.

Na fase atual (2.^a) os objetivos passam por:

- Alargar o modelo de dados à definição e tratamento de tarefas;
- Criação de vistas baseadas em funções para *stakeholders* em múltiplos projetos;
- Apoio de repositórios e sistemas de identificação para ligação com sistemas de armazenamento de dados ou arquivos de longo prazo;

¹⁰ RDMO: First projekt phase. Disponível em: <https://rdmorganiser.github.io/en/first-project-phase/>

- Apoio à apresentação de sugestões de repositórios de dados adequados e, eventualmente, estimativas de custos com base nos respetivos termos;
- Extensão das interfaces aos sistemas de autenticação e autorização;
- Participação no esforço de criação de um formato de intercâmbio internacional para o PGD;
- Melhoria dos procedimentos técnicos de instalação, manutenção e atualização;
- Sustentabilidade e construção da comunidade;
- Divulgação e formação.

Informação do Software

Open Source disponível em: <https://github.com/rdmorganiser/rdmo>

Documentação essencial à sua instalação e manutenção:

<https://rdmo.readthedocs.io/en/latest/index.html>

Termos de uso: <https://rdmorganiser.github.io/en/protection/>

Comunidade de utilizadores

Existe uma lista de discussão geral do RDMO: rdmo@listserv.dfn.de

De igual forma, o utilizador / interessado tem a possibilidade de fazer perguntas diretamente à equipa de trabalho: rdmo-team@listserv.dfn.de

Têm implementado reuniões abertas à comunidade:

https://rdmorganiser.github.io/workshop022020_programm/

Possuem canal de comunicação no Slack: rdmo.slack.com

Modelos para Financiadores

Permite o trabalho colaborativo e a extração dos dados em diferentes formatos, de acordo com os requisitos dos financiadores.

Casos de uso

Os casos de uso desta ferramenta são feitos com base nas experiências das instalações já em curso ou em fase de teste um pouco por toda a Alemanha:

https://rdmorganiser.github.io/workshop022020_programm/

Informação de suporte

Informação sobre instalação, desenvolvimento, configuração, gestão e atualizações:

<https://rdmo.readthedocs.io/en/latest/index.html>

Neste momento estão a desenvolver melhoramentos e a adicionar interfaces de comunicação com infraestruturas institucionais, como repositórios. Alguma da [documentação¹¹](#) (tutoriais, materiais de apoio, documentação, material de divulgação) já existe, mas maioritariamente em língua alemã. Prevê-se a tradução para outros idiomas.

Planos de Custos

Custos para o cliente (estimativa):

- Uma vez para a instalação: 5.000 euros
- Por ano (se necessário, proporcionalmente no ano de início): 3.000 euros
- Os preços podem cair em perspetiva, sendo (ainda) difícil de calcular.

DMPtool

	<p>Designação DMPTool https://dmptool.org</p> <p>Proprietários University of California Curation Center (UC3), DataONE, Digital Curation Centre (DCC), Smithsonian Institution, University of California, Los Angeles Library, University of California, San Diego Libraries, University of Illinois, Urbana-Champaign Library e University of Virginia Library.</p>
<p>Breve descrição do serviço</p> <p>O DMPTool é uma ferramenta <i>online</i> que permite criar e editar planos de gestão de dados personalizados. Esta ferramenta permite: criar planos de gestão de dados prontos a usar para financiadores, cumprir os respetivos requisitos e obter orientações para os dados do plano de gestão.</p>	

Informação do Software

Software *Open Source* que faz parte do ROADmap, à semelhança do DMPonline.

Comunidade de utilizadores

A comunidade que utiliza o DMPTool é bastante alargada. A ferramenta está integrada no GitHub: <https://github.com/topics/dmptool>.

¹¹ RDMO: Documentation. Disponível em: <https://rdmorganiser.github.io/en/documentation/>

Modelos para Financiadores

Esta ferramenta disponibiliza modelos para financiadores, permitindo a criação de novos modelos de acordo com os requisitos da organização ou financiador.

Casos de uso

- Cerca de 47.300 utilizadores;
- 268 organizações;
- Total de 44.400 planos de gestão de dados (informação recolhida a 19/05/2020).

Informação de suporte

Do ponto de vista de suporte, as orientações são providenciadas através do GitHub e incluem:

- Orientações sobre os planos de gestão de dados;
- Criação e customização de modelos.

Planos de Custos

Informação não disponível publicamente.

ezDMP - NSF

	<p>Designação ezDMP https://ezdmp.org</p> <p>Proprietários National Science Foundation MIT License Copyright 2018 ezDMP</p>
<p>Breve descrição do serviço</p> <p>Trata-se de uma instalação do DMPTool adaptado às necessidades da National Science Foundation (NSF), que criou diretrizes específicas para a produção destes planos para facilitar os investigadores no cumprimento dos seus requisitos. O ezDMP foi desenvolvido para garantir que os investigadores forneçam tudo o que a NSF requer, de forma clara e organizada, dando igualmente a possibilidade de recolha de informação sobre as análises significativas no que diz respeito à conformidade dos dados.</p> <p>O ezDMP evoluiu de uma primeira abordagem feita pela Interdisciplinary Earth Data Alliance - IEDA DMPTool (https://www.iedadata.org/dmp/), desenvolvida inicialmente para satisfazer as necessidades da sua comunidade, na recolha dados terrestres provenientes dos ambientes marinho, terrestre e polar, no contexto de projetos financiados pela NSF.</p>	

Informação do Software

Baseado na ferramenta DMPTool:

GitHub: <https://github.com/ezdmp/ezDMP-Site>

Comunidade de utilizadores

Não tem qualquer informação associada.

Modelos para Financiadores

Disponibiliza modelo de PGD para a National Science Foundation (NSF).

Casos de uso

As melhorias na ferramenta foram feitas com base nos contributos dos utilizadores, através da realização de inquérito.

*Open Access to Research Artifacts: Implementing the Next Generation Data Management Plan*¹².

Informação de suporte

A informação disponibilizada é de carácter generalista enquanto organismo financiador, no contexto da gestão de dados de investigação:

- **Exemplo 1:** <https://mcbblog.nsfbio.com/2019/07/18/data-management-plans-tips-from-dcl-19-069/>
- **Exemplo 2:** <https://www.nsf.gov/pubs/2019/nsf19022/nsf19022.pdf>

Exemplo de um PGD extraído a partir do ezDMP:


<https://zenodo.org/record/3247756#.XsaX2GhKhPZ>

Planos de Custos

Informação não disponível publicamente.

¹² Stodden, V., Ferrini, V., Gabanyi, M., Lehnert, K., Morton, J., & Berman, H. (2019). Open access to research artifacts: Implementing the next generation data management plan. *Proceedings of the Association for Information Science and Technology*, 56(1), 481-485. <https://web.stanford.edu/~vcs/papers/STODDEN-AM19ASIST.pdf>

easyDMP

	<p>Designação easyDMP https://easydmp.eudat.eu</p> <p>Proprietário EUDAT</p>
<p>Breve descrição do serviço</p> <p>O easyDMP (https://www.eudat.eu/catalogue/easyDMP), desenvolvido pelo Sigma2 em colaboração com o EUDAT, tem como objetivo disponibilizar uma ferramenta que permita aos investigadores com o mínimo de experiência em gestão de dados, uma forma simples de criar planos de gestão de dados. Para tal, esta ferramenta procede à transformação das orientações das agências de financiamento para a gestão de dados, num conjunto de questões de fácil resposta, muitas delas de resposta pré-definida.</p> <p>Este serviço simplifica a tarefa de criação de planos de gestão de dados. O processo de criação começa pela seleção do modelo de PGD e de seguida a resposta a um conjunto de perguntas de fácil resposta que abrangem todos os aspetos das normas das agências de financiamento. Permite ainda a partilha dos PGDs com outros utilizadores, bem como a sua publicação e exportação. Os planos podem ser atualizados durante o período de desenvolvimento do projeto ao qual está associado, a fim de refletir com precisão o processo de gestão dos dados.</p> <p>Características:</p> <ul style="list-style-type: none"> • Autenticação: B2Access, Dataporten (Feide federated identity), Facebook, Google, LinkedIn; • PGD <i>machine-readable</i>. 	

Informação do Software

Open Source.

Comunidade de utilizadores

GitHub: <https://github.com/hmpf/easydmp/>

Documentação para o utilizador: <https://www.sigma2.no/easydmp-user-documentation>

Serviço utilizado a nível nacional na Noruega (*Norwegian Research Council*).

Modelos para Financiadores

Disponibiliza os seguintes modelos de PGDs:

- Modelo Science Europe (<https://www.sigma2.no/science-europe-template-practical-guide>);
- Modelo Horizonte 2020.

Possibilita também a criação de novos à medida das necessidades, como por exemplo modelos específicos para instituições e comunidades de investigação. Para tal, disponibiliza um interface de administrador a partir do qual é possível criar novos modelos ou atualizar os existentes. Para saber mais contactar sigma2@uninett.no.

Casos de uso

573 utilizadores, 597 planos, 156 organizações (informação recolhida a 22/05/2020).

Informação de suporte

Guia de iniciação: <https://www.sigma2.no/data-planning#get-access>

Guia do utilizador: <https://www.sigma2.no/easydmp-user-documentation>

Suporte: support@easydmp.sigma2.no

Planos de Custos

O serviço é gratuito e acessível a todos.

Política de utilização: <http://hdl.handle.net/11304/e43b2e3f-83c5-4e3f-b8b7-18d38d37a6cd>

ARIADNE

	<p>Designação DMP Researcher Template for Archaeological Datasets https://vast-lab.org/dmp/ariadneplus/form/</p> <p>Proprietário ARIADNEplus / Parthenos</p>
<p>Breve descrição do serviço</p>	
<p>O ARIADNEplus é um projeto financiado pela Comissão Europeia, sob o programa Horizonte 2020.</p> <p>Esta iniciativa disponibiliza diversos serviços, de entre os quais se destaca um plano de gestão de dados online denominado de DMP Researcher Template for Archaeological Dataset, dirigido a arqueologistas e investigadores nesta área.</p>	

Informação do Software

Open Source.

Comunidade de utilizadores

A comunidade de utilizadores inclui as associações arqueológicas europeias mais relevantes, investigadores na área de arqueologia, agências de património nacional e profissionais na área do património.

Modelos para Financiadores

A ferramenta inclui um modelo geral de plano de gestão de dados que cumpre os requisitos do programa Horizonte 2020 em acesso aberto, bem como os princípios FAIR específicos à área da arqueologia.

Casos de uso

O DMP Researcher Template for Archaeological Dataset centra-se especificamente no domínio da arqueologia.

Informação de suporte

A ARIADNEplus disponibiliza sessões de formação e workshops gerais.

Não foi encontrada nenhuma informação específica referente à aplicação de planos de gestão de dados.





Planos de Custos





Informação não disponível publicamente.

Análise comparativa de funcionalidades, requisitos e custos

No quadro de uma análise comparativa mais detalhada, optou-se por destacar quatro ferramentas. A seleção destas ferramentas resulta da robustez técnica apresentada, do universo de utilizadores e da sustentabilidade do sistema no contexto de serviços de âmbito nacional ou no quadro de infraestruturas europeias de investigação, sendo também ponderado o potencial de adoção aos cenários de aplicação em Portugal. Neste sentido, a análise comparativa das ferramentas é desenvolvida com base em catorze características e funcionalidades de cariz eminentemente técnico e administrativo, identificando se cada uma das quatro ferramentas estudadas possui essas funcionalidades. São analisadas por exemplo, software *open source* vs proprietário ou disponibilidade de APIs, bem como as funcionalidades relacionadas com conteúdo ou utilização, avaliando por exemplo a comunidade de utilizadores e casos de uso das ferramentas, ou se os sistemas possuem meios para criar modelos de planos ou informação adaptada a cada financiador.

Análise de funcionalidades e características

Aplicações				
Open source vs. Proprietário	Open source	Open source	Open source	Open source
Serviços Cloud vs Local	Disponibiliza os dois serviços	Disponibiliza os dois serviços	Disponibiliza os dois serviços	Disponibiliza os dois serviços
Utilizadores e Casos de Uso	<p>A comunidade que utiliza o DMPonline é bastante alargada:</p> <ul style="list-style-type: none"> - Cerca de 17.600 utilizadores; - 203 organizações; - Presente em 89 países; - Total de 23.000 planos de gestão de dados (informação recolhida a 19/05/2020). <p>Casos de uso: O DMPonline dá apoio e suporte a várias instituições universitárias britânicas através do <i>Higher Education Funding Council of England</i></p>	<p>A comunidade que utiliza o ARGOS é bastante alargada:</p> <ul style="list-style-type: none"> - 132 organizações; - 85 PGDs; - 91 descrições de datasets (informação recolhida a 26/05/2020). <p>Casos de uso: Foi apresentado pela CE como um dos serviços disponíveis para suporte à gestão de dados de investigação em torno do COVID-19. A FECYT utiliza o ARGOS. O serviço encontra-se também em disseminação por vários países Europeus, pela rede de secretariados nacionais do OpenAIRE</p>	<p>As instituições que utilizam o Data Stewardship Wizard estão maioritariamente relacionadas com o domínio de Ciências da Vida.</p> <p>Casos de uso: Experiências das instalações já em curso</p>	<p>Vários canais de comunicação com a comunidade.</p> <p>Casos de uso: Experiências das instalações já em curso ou em fase de teste em vários institutos alemães</p>

Aplicações				
Software open source (especificações de instalação)	O software é uma aplicação Ruby on Rails. Terá de estar instalado o Ruby 1.9.2 ou versão superior no servidor. Terá de estar também instalado o MySQL V5.0 ou versão superior https://github.com/CottageLabs/DMPOnline/wiki/1.-Local-Installation	OpenDMP GitLab https://gitlab.eudat.eu/dmp/OpenAIRE-EUDAT-DMP-service-pilot	Aplicação integrada num servidor da CESNET Ou Via Docker https://docs.ds-wizard.org/en/latest/admin/installation.html	A instalação difere de acordo com o sistema operativo. Consultar o seguinte link para mais informações: https://rdmo.readthedocs.io/en/latest/installation/prerequisites.html
Ferramentas Actionable / Integração	<i>Machine-actionable</i>	<i>Machine-actionable</i>	<i>Machine-actionable</i>	<i>Machine-actionable</i>
Autenticação	Autenticação requerida	Não é necessária autenticação. O login poderá ser efetuado via OpenAIRE e EUDAT AAI, ORCID, Google e contas de redes sociais (Facebook, Twitter)	Autenticação interna via ELIXIR AAI e Google. Autenticação externa por OpenID	Autenticação requerida. Três possibilidades: django-allauth, servidor LDAP ou Shibboleth. https://rdmo.readthedocs.io/en/latest/configuration/authentication/
APIs (Disponibilidade de webservices, integração de APIs externas)	Informação em: https://github.com/DMPRoadmap/roadmap/wiki/API-Documentation	Informação em: http://develop.openaire.eu/	Informação em: https://github.com/ds-wizard/docs/blob/master/docs/user/integrations.rst	Informação em: https://rdmo.readthedocs.io/en/latest/administration/api.html
Inserção de DOI	Pode ser atribuído um DOI ao PGD	Pode ser atribuído um DOI ao PGD	Pode ser atribuído um DOI ao PGD	Pode ser atribuído um DOI ao PGD
Versionamento	N/A (podem ser efetuadas edições, no entanto não existe controle de versionamento)	Sim (integrado no Zenodo)	Sem informação	Sim
Opções de exportação dos planos	PDF, DOC, XML	PDF, DOC, XML, JSON	PDF, DOC, LaTeX, JSON	XML, DOCX, ODT





Aplicações				
Opção de publicação online DOI	Sim (website da aplicação)	Sim (Zenodo)	Sim (aplicação)	Sim (aplicação)
Guidelines (geral ou domínio específico)	Geral	Geral	Geral	Geral
Modelos	Disponibiliza modelos de financiadores, podendo haver customização	Disponibiliza modelos de financiadores, podendo haver customização	Disponibiliza modelos de financiadores, podendo haver customização	Disponibiliza modelos de financiadores, podendo haver customização
Orientação adaptada a instâncias (financiador /instituições)	Sim	Sim	Sim	Sim





Identificação e validação de requisitos base

No quadro do desenvolvimento de um serviço nacional de apoio à elaboração de PGDs foram identificados os seguintes requisitos de referência:

- Software *Open Source*;
- Customização de modelos (adaptação a serviço guiado pela política da FCT);
- Integração com autenticação centralizada;
- Integração com PTCRIS;
- Normativo PTCRIS:
 - Modelo de dados (CERIF);
 - Ontologias;
 - Identificadores únicos (autor);
 - Identificadores organizacionais;
 - Dicionário de interoperabilidade.
- Interoperabilidade com outros sistemas;
- *Machine-actionable*;
- Comunicação e disseminação do serviço;
- Informação de planeamento e tarefas de gestão de dados relevantes, ao longo de todo o ciclo de vida dos dados de investigação.

Apresenta-se na seguinte tabela a validação da conformidade de cada requisito base para o conjunto das quatro ferramentas analisadas, tendo por base a informação previamente recolhida diretamente dos responsáveis por cada serviço.

APLICAÇÕES					
Software Open Source		✓	✓	✓	✓
Customização de modelos e adaptação a serviço nacional, guiado pela política da FCT		✓	✓	✓	✓
Integração com autenticação centralizada		✓	Não é necessária autenticação. Definições ao nível do login.	✓	✓
Integração com PTCRIS		Apresenta uma ligação ao API do OpenAIRE com a recolha de informação de <i>grants</i> da CE (programa H2020)	✓ Informação no OpenAIRE (todos os <i>grants</i> incluindo os da FCT)	-	-
Normativo PTCRIS	Modelo de dados (CERIF)	Ligação externa inclui standard CERIF (via RESTful API) / SWORD2	Ligação externa inclui CERIF	-	Ligação externa inclui CERIF
	Ontologias	Como maDMP considera tecnologia semântica	Como maDMP considera tecnologia semântica, no âmbito dos serviços do OpenAIRE e EOSC	Utiliza serviços do ELIXIR - Ontologia EDAM	-
	Identificadores únicos (autor)	Integração com ORCID	Integração com ORCID	Integração com ORCID	Integração com ORCID




APLICAÇÕES					
	Identificadores organizacionais	Integração com ROR (que por sua vez mapeia com ISNI)	OpenAIRE integra com ISNI, OpenDOAR, re3data	Ainda não existe integração, mas poderá ser adicionada	A nova versão do software permite a integração de APIs externas. Podem ser criados <i>plugins</i> para importação e exportação de informação de projetos
	Dicionário de interoperabilidade	Mapeamento entre os modelos e os conceitos definidos no CASRAI	OpenAIRE está alinhado com os termos de publicações CASRAI	Ligação entre o modelo DS-Wizard e a taxonomia CASRAI	Sem informação
Interoperabilidade com outros sistemas		✓	✓	✓ Mais limitado	✓
Machine-actionable		✓	✓	✓	✓
Comunicação e disseminação do serviço		✓ Suporte elevado à comunidade de utilizadores	✓ Suporte elevado à comunidade de utilizadores	✓ Suporte mais centrado no domínio das Ciências da Vida e no quadro do FAIRsharing	✓ Suporte à comunidade de utilizadores

Informação sobre custos

Verificámos que a informação disponível online sobre planos de custos para cada uma das ferramentas é limitada. Como tal, efetuámos um contacto com o serviço de apoio ao cliente de cada uma das aplicações.

De forma a receber a cotação final, teremos de transmitir a informação sobre os requisitos finais para a implementação do serviço nacional.

A tabela seguinte sumariza a informação sobre os custos de cada uma das aplicações.

			
<p>Dois níveis de assinatura: um pacote básico e um pacote melhorado¹³</p> <p>Em termos do modelo de assinatura existem assinaturas anuais, em que todos os anos o cliente será informado sobre o preço para o ano seguinte ou uma oferta de uma assinatura de 3 anos, em que a taxa do ano corrente será aplicada e serão atribuídos 10% de desconto¹⁴</p> <p>Estimativa de custos (cloud):</p> <ul style="list-style-type: none"> - Pacote básico: 6.000 £ - Pacote melhorado: 7.900 £ 	<p>Os custos são neste momento cobertos pelo projeto OpenAIRE Advance a finalizar em fevereiro de 2021</p>	<p>Atualmente não há qualquer custo em relação a nenhum dos planos oferecidos. De acordo com a informação obtida, o custo está a ser neste momento coberto pelo nó ELIXIR CZ. É possível que no futuro esta situação se altere</p>	<p>Estimativa de custos:</p> <ul style="list-style-type: none"> - Instalação: 5.000 € - Valor anual: 3.000 €

¹³ Com o pacote básico é possível criar uma orientação institucional, criar modelos próprios, comentar e dar feedback sobre os planos, e ver as estatísticas dos utilizadores. O benefício da assinatura melhorada corresponde à possibilidade de personalizar o front-end da ferramenta para que se pareça com um serviço institucional.

¹⁴ O pagamento pode ser dividido em três pagamentos anuais, pelo que a cada ano de subscrição apenas será enviada uma fatura.

ESTUDOS, INICIATIVAS E PROJETOS EM PORTUGAL SOBRE PGDs

A nível nacional, a consciencialização desta área essencial do processo de investigação já começa a ter alguma expressão, quer no desenvolvimento de estratégias de estudo e investigação na área, quer ao nível da sua promoção e discussão na comunidade, com fim último de criar e capacitar recursos. A este nível é de ressaltar a importância do Fórum de Gestão de Dados de Investigação (GDI) (<https://forumgdi.rcaap.pt/>) que, desde 2016, trabalha para a criação e consolidação de uma comunidade em torno da GDI.

Neste contexto, destacamos o workshop “[Planos de Gestão de Dados: ferramentas e requisitos dos financiadores](#)”¹⁵, apresentado no 1.º Fórum e onde, em representação dos projetos [OpenAIRE](#) (www.openaire.eu) e [FOSTER](#) (www.fosteropenscience.eu), foram enunciados princípios, requisitos da CE para a GDI, assim como também se explorou a ferramenta - [DMP Online](#) (<https://dmponline.dcc.ac.uk/>) - para criação e manutenção de planos de gestão de dados. Outros recursos online, experiências de várias instituições também foram disponibilizados, como por exemplo de: [ANDS](#) (<https://www.andis.org.au/>), [DCC](#) (<https://www.dcc.ac.uk/>), [RDNL](#) (<https://researchdata.nl/>) e igualmente vários materiais de formação produzidos pelos projetos já referidos.

No que diz respeito à capacitação de profissionais, que desenvolvem a sua atividade nesta área, bem como à comunidade de investigadores, têm vindo a ser disseminados webinars, com periodicidades anual, promovidos no âmbito dos projetos FOSTER e OpenAIRE, uns de carácter mais genérico, no que diz respeito ao [cumprimento dos requisitos para a gestão de dados, no contexto do H2020](#)¹⁶, outros particularizando a [elaboração de planos de gestão de dados de investigação](#)¹⁷.

Outras experiências de aproximação à comunidade de investigadores, por forma a atuar como facilitador, nesta fase de transição e adaptação aos novos procedimentos, têm vindo a ser partilhadas, também neste [Fórum GDI](#) (<https://forumgdi.rcaap.pt/>). Assim, destacamos um método desenvolvido pela [BioData.pt](#) (<https://biodata.pt/>) (infraestrutura nacional para dados biológicos e representante português do [ELIXIR](#) (<https://elixir-europe.org/>)), no âmbito do

¹⁵ http://confdados.rcaap.pt/wp-content/uploads/2016/09/ForumGDI_workshop_PedroPrincipe.pdf

¹⁶ Webinar OpenAIRE Portugal: Requisitos para Gestão de Dados Abertos no H2020. Disponível em: <https://www.youtube.com/watch?v=wP8AqMob0lc&t=316s>

¹⁷ Webinar "Planos de Gestão de Dados". Disponível em: <https://www.youtube.com/watch?v=-VvktOA3FCI&t=249s>

programa de treino [Ready for Bio-Data Management](#)¹⁸. [Este método](#)¹⁹ consiste na criação de um PGD integrando uma sequência de 5 passos de recolha de informação, de acordo com o RDA DMP *common standard*:

1. Dados administrativos;
2. Projeto e financiamento;
3. Caracterização do PGD;
4. Caracterização dos dados;
5. Preservação e publicação.

Esta informação é preenchida em grupo num “PGD Canvas”, com base no *Business Model Canvas*, ou alternativamente, na versão de treino à distância, num documento *Google Spreadsheets* com a mesma estrutura.

Outro exemplo decorreu no âmbito do desenvolvimento do [projeto TAIL](#) (<https://www.inesctec.pt/pt/projetos/tail>) (INESC TEC), que desenvolveu um [método colaborativo de criação e construção de um PGD](#)²⁰, por forma a amenizar os custos de esforço e tempo requeridos para a elaboração de um plano pormenorizado. Neste processo foram incluídas entrevistas aos investigadores, análise das suas publicações e dos dados de investigação relacionados com o projeto, bem como PGDs análogos na área. Após esta primeira fase, os PGDs seriam apresentados aos investigadores para aperfeiçoamento e eventuais correções, com o propósito de alcançar uma versão final.

Por fim, dar conta do desenvolvimento do [MOOC “Essencial da Gestão de Dados de Investigação”](#)²¹ que integra o portfólio da plataforma NAU (<https://www.nau.edu.pt/>). Este curso, além de se centrar nos principais aspetos associados à GDI, ao longo do ciclo de vida de investigação, dedica em exclusivo um módulo aos PGDs, promovendo a realização de atividades práticas, com recurso a várias tipologias e suportes de informação, tornando-o bastante dinâmico e interativo. Quanto ao seu público alvo este destina-se, essencialmente, a investigadores e alunos de doutoramento, bem como outros profissionais que exerçam atividades de suporte à gestão de dados de investigação.

¹⁸ Ready 4 Bio Data Management: <http://ready4biodatamanagement.biodata.pt/index.html>

¹⁹ Cardoso, J., Faria, D., Melo, A., & Borbinha, J. (2019). Planos de Gestão de Dados: Um novo método de treino. 5.º Fórum GDI. https://forumgdi.rcaap.pt/wp-content/uploads/2019/11/1-JoaoCardoso_5ForumGDI.pdf

²⁰ Karimova, Y., & Ribeiro, C. (2019). O método colaborativo entre curadores e investigadores na preparação de um DMP e PIA. 5.º Fórum GDI. https://forumgdi.rcaap.pt/wp-content/uploads/2019/11/2-YuliaKarimova_5ForumGDI.pdf

²¹ MOOC “Essencial da Gestão de Dados de Investigação”. Disponível em: https://lms.nau.edu.pt/courses/course-v1:SEC-GERAL.MEC+GDI101+2019_T1/about


Tabela resumo de algumas das principais iniciativas apresentadas publicamente:

Fórum GDI - workshop sobre planos de gestão de dados (2016)	+ info
Ready for BioData Management (2019-2020)	+ info
OpenAIRE webinars (2018 e 2019)	+ info
Projeto TAIL - DMPs (2019)	+ info
MOOC “Essencial GDI” (2019-2020)	+ info

SERVIÇOS NACIONAIS PARA PGDs

Nesta seção do relatório pretendem-se apresentar iniciativas nacionais de serviços de apoio à gestão de dados de investigação, que incluem ferramentas de elaboração de PGDs no seu portfólio de serviços para a gestão de dados. Apresentam-se de forma breve, seis casos de estudo, nomeadamente de Espanha, Bélgica, França, Noruega, Dinamarca e Alemanha que oferecem uma aplicação para elaboração de PGDs disponibilizada a nível nacional ou regional.

PaGoDa - Espanha


 PGDOnline	DMPonline http://pgd.consorciomadrono.es
<p>O sistema PaGoDa, é a aplicação para elaboração de planos de gestão de dados disponibilizada pelo Consortio Madroño, no âmbito do projeto de serviços de apoio à gestão de dados - e-cienciaDatos.</p>	

O [Consortio Madroño](http://www.consorciomadrono.es/investigam/pagoda/) (www.consorciomadrono.es/investigam/pagoda/) disponibiliza, desde 2014, uma ferramenta online designada de PGDOnline, com adaptação e tradução para espanhol da ferramenta de gestão DMPonline desenvolvida pelo Digital Curation Centre no Reino Unido, tendo em conta as orientações para a gestão de dados incluídas no programa Horizonte 2020.

As bibliotecas e instituições académicas integradas no Consorcio Madroño constituem os principais utilizadores desta ferramenta.

O sistema PaGoDa oferece igualmente materiais de apoio e um serviço de consulta aos seus membros.

dmponline.be - Bélgica

	<p>DMPonline https://dmponline.be</p>
<p>Consórcio nacional Belga iniciado no ano de 2017, por um conjunto de Universidades e Instituições do Ensino Superior, contando atualmente com 12 parceiros, para os quais é disponibilizado o serviço de planeamento de gestão de dados baseado no DMPonline.</p>	

O serviço dmponline.be para a criação de Planos de Gestão de Dados para as Universidades e Instituições de Ensino Superior da Bélgica, utiliza uma instância nacional do DMPonline (versão 4.1).

A iniciativa da criação deste serviço teve início em 2015 com a instalação de uma instância local pela Biblioteca da Universidade de Ghent. A configuração desta instalação foi customizada de acordo com as necessidades da instituição e serviu de piloto para aquilo que viria a ser a versão coletiva da ferramenta para as Instituições de Ensino e de Investigação Belgas.

Em janeiro de 2017 foi fundado o Consórcio *DMPbelgium*, por 7 instituições de Ensino Superior e de Investigação, disponibilizando uma ferramenta de planeamento de gestão de dados para os utilizadores das instituições aderentes. Em 2018 o consórcio viu o seu número de participantes aumentar, sendo atualmente 12 parceiros no total.


O DMPonline.be disponibiliza um conjunto de modelos de PGDs, tanto ao nível de financiadores no contexto nacional, como também modelos institucionais (e subgrupos como escolas e unidades de investigação), acompanhados por um conjunto de guias de suporte.

O código fonte é disponibilizado no GitHub em:

https://github.com/DMPbelgium/DMPonline_v4.

Quanto ao registo por parte dos utilizadores, o serviço oferece um sistema de autenticação federada (Belnet Federation), e via ORCID.

DMP OPIDoR - França

	DMPRoadmap https://dmp.opidor.fr
<p>O DMP OPIDoR corresponde à instância francesa do sistema nacional de planos de gestão de dados e, como tal, é direcionada aos investigadores franceses. Esta aplicação é disponibilizada pelo Centre Nationale de la Recherche Scientifique (CNRS).</p>	


O DMP OPIDoR é uma ferramenta *Open Source* e baseia-se no DMPRoadmap, que é desenvolvido conjuntamente entre o Digital Curation Centre (DCC) e a University of California Curation Center (UC3). O código está disponível no GitHub.

A aplicação inclui vários modelos que representam os requisitos de diferentes financiadores e instituições. Os utilizadores são convidados a responder a três questões, de modo a que seja apresentado o modelo mais apropriado. Este processo é suportado por orientações fornecidas por um conjunto de financiadores, universidades e domínios de investigação.

O DMP OPIDoR pode ser customizado de acordo com as instituições e comunidades académicas. Na realidade, podem ser adicionados modelos para utilizadores e orientações personalizadas de acordo com o suporte e serviços locais. Podem também ser adicionados exemplos para ajudar os utilizadores a criarem o seu PGD.

A aplicação disponibiliza uma lista dos PGDs públicos em: https://dmp.opidor.fr/public_plans.

DMP Norway - Noruega (NSD)

	NSD https://dmp.nsd.no
<p>Serviço nacional disponibilizado pelo arquivo e centro de dados de pesquisa norueguês.</p>	

O Norwegian Centre for Research Data (NSD) constitui um arquivo e centro nacional de dados de investigação. Esta instituição visa garantir o acesso aberto a dados de investigação, melhorando as oportunidades de pesquisa empírica através de uma ampla gama de serviços de informação e suporte. O NSD disponibiliza um serviço de planos de gestão de dados designado de NSD DMP.


O NSD DMP cumpre com os requisitos do Conselho de Investigação Norueguês e da Science Europe. Apresenta a funcionalidade *machine-actionable* e inclui um módulo que classifica os dados em termos de confidencialidade e nível de segurança.

O NSD DMP inclui um guia que auxilia os utilizadores a selecionarem os arquivos e repositórios nacionais mais adequados para os seus dados. Este guia utiliza APIs da re3data.org.

A informação da aplicação está disponível no GitHub:

<https://github.com/makinin/rdbhapi>.

DMP Online - Dinamarca (DeiC)

	<p>DMPRoadmap https://dmponline.deic.dk</p>
<p>Serviço nacional disponibilizado pelo DeiC (Danish e-infrastructure Cooperation) a todas as instituições de investigação e ensino superior Dinamarquesas, tendo por base a utilização do DMPRoadmap.</p>	

O serviço de PGDs trata-se de uma instalação nacional do DMPOnline.

O serviço é disponibilizado pelo DeiC (Danish e-infrastructure Cooperation), e administrado conjuntamente pela Royal Danish Library e Technical University of Denmark (DTU Library), estando disponível para todos os colaboradores das instituições de investigação dinamarquesas.

O DeiC tem como função a coordenação da infraestrutura digital dinamarquesa para as 8 universidades nacionais, assegurando a disponibilização de serviços de computação, armazenamento e infraestrutura de rede para a investigação nacional, ensino e inovação. Adicionalmente, assume também a responsabilidade para disponibilização de serviços na área da gestão e armazenamento de dados de investigação, tais como: serviço para a elaboração de planos de gestão de dados (DMP Online), serviços de armazenamento (Sciencedata.dk) e partilha de dados (FileSender), serviço de atribuição de identificadores persistentes (DataCite) e serviço de suporte e desenvolvimento de diretrizes para a gestão e armazenamento de dados.

Quanto ao software para o serviço DMP Online, este é baseado no código base *open source* DMPRoadmap, que é desenvolvido conjuntamente pelo DCC e UC3. A gestão do serviço é assegurada pela DTU Library e Royal Danish Library. O DMPOnline pode ser customizado e disponibiliza guias e modelos de diferentes agências de financiamento.

Para que os utilizadores possam utilizar o serviço, devem criar uma conta de utilizador e de seguida poderão opcionalmente associar a sua conta com a sua instituição, utilizando o serviço de autenticação WAYF.

A sua utilização é gratuita para instituições de investigação e de ensino superior Dinamarquesas. Quanto aos custos de operação, os mesmos são assegurados pelo orçamento geral do DeIC.


O código fonte é disponibilizado no GitHub em:

<https://github.com/DMPRoadmap/roadmap>

Termos do serviço:

www.deic.dk/sites/default/files/uploads/PDF/DMPonline_terms_v181210.pdf

RDMO - Alemanha

	Research Data Management Organiser https://rdmorganiser.github.io
<p>O RDMO está acessível quer às instituições, bem como aos investigadores, para o desenvolvimento e planeamento, de forma autónoma da gestão dos seus dados de investigação.</p>	

Conforme já referido na apresentação desta ferramenta (terceira parte deste relatório), o RDMO está acessível quer às instituições, bem como aos investigadores, para o desenvolvimento e planeamento, de forma autónoma da gestão dos seus dados de investigação. Esta ferramenta possibilita a congregação de toda a informação de planeamento e tarefas de gestão de dados relevantes, ao longo de todo o ciclo de vida dos dados de investigação.

O âmbito da sua aplicação é ajustável, quer se trate de projetos de maior ou menor dimensão, tendo eventualmente já sido alargada a outras organizações parceiras [AIP](https://www.aip.de/en/institute/) (<https://www.aip.de/en/institute/>), [FHP](https://www.en.fh-potsdam.de/) (<https://www.en.fh-potsdam.de/>) e [KIT Library](http://www.kit.edu/english/index.php) (<http://www.kit.edu/english/index.php>) que estão a colaborar, com os utilizadores do RDMO para melhorar a sua utilização.

A sua implementação está a ser reforçada (funções e interfaces) ao nível das infraestruturas institucionais, como é caso de repositórios, sistemas de bilhética e infraestruturas de autenticação.

CONCLUSÕES

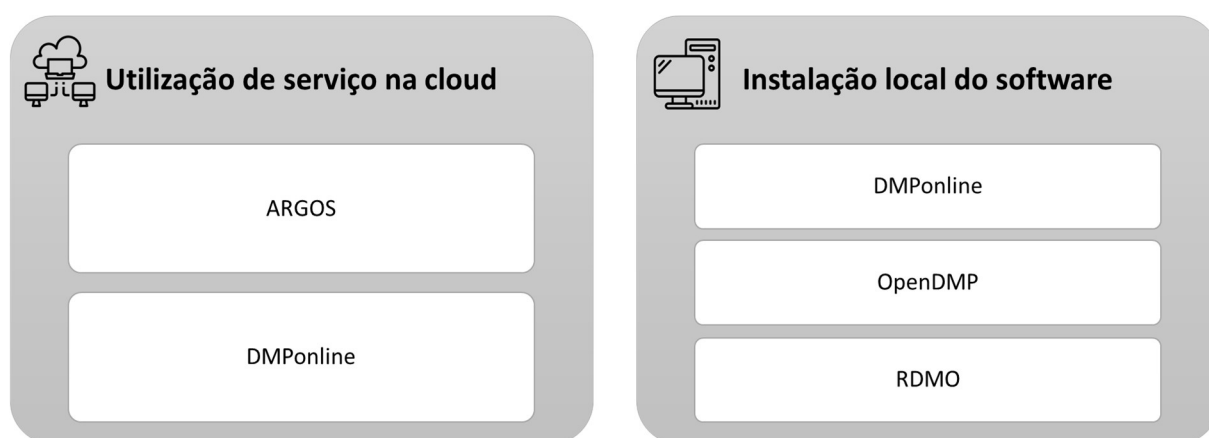
O presente relatório identificou as principais ferramentas de elaboração de planos de gestão de dados de investigação e estabeleceu uma análise comparativa de funcionalidades e características, tendo em vista a disponibilização de um serviço nacional para apoiar a elaboração de PGDs.

Das nove soluções de aplicações ou ferramentas analisadas foram identificadas quatro ferramentas – **DMPonline**, **ARGOS/OpenDMP**, **Data Stewardship Wizard**, **RDMO** (Research Data Management Organiser) – que devem ser consideradas para o desenvolvimento de um serviço nacional a desenvolver em Portugal.

Estas quatro ferramentas possuem características que lhes conferem robustez técnica e garantia de sustentabilidade a médio-longo prazo, apresentando-se ainda casos de uso e um universo de utilizadores que asseguram tratar-se de soluções adequadas para aplicação a serviço à escala nacional.

As quatro ferramentas selecionadas foram alvo de um estudo mais detalhado realizado a partir da análise de catorze características ou funcionalidades. Esta análise comparativa revelou com maior clareza as recomendações técnicas e tipologia de soluções a implementar. Assim, as soluções preconizadas com base no presente estudo estão estruturadas em dois eixos de intervenção:

- 1º) Utilização de serviço na cloud - recurso à adesão ao serviço online,
- 2º) Instalação local do software – recurso à instalação e configuração de uma instância local.



Tipos de soluções propostas e ferramentas selecionadas

1º) Utilização de serviço na cloud - recurso à adesão ao serviço online e utilização da aplicação genérica de PGDs..

- Esta solução de utilização livre do serviço universal implicará a parametrização de modelos de planos para o financiador nacional, assim como o acesso a uma área de administração de modelos, informação de suporte e de administração de utilizadores e ainda assegurar a tradução do sistema para português.
- Para esta solução propõe-se duas ferramentas que possuem as características essenciais e as funcionalidades básicas para o estabelecimento de um serviço à escala nacional:
 - ARGOS (<https://argos.openaire.eu/>) – OpenAIRE (EU)
 - DMPonline (<https://dmponline.dcc.ac.uk/>) – Digital Curation Centre (UK)

2º) Instalação local do software – recurso à instalação e configuração de uma instância local do software de PGDs.

- Esta solução implica a instalação da versão mais recente do software e respetiva parametrização, que inclui tradução e preparação de modelos de planos para o financiador nacional, e a adaptação às especificações técnicas no quadro de integrações de serviços nacionais, nomeadamente de autenticação ou identificadores.
- Para esta solução propõem-se dois software *Open Source* que possuem as características fundamentais e cumprem integralmente os requisitos funcionais para o estabelecimento de um serviço nacional de apoio à elaboração de PGDs:
 - DMPonline (<https://dmponline.dcc.ac.uk/>) – Digital Curation Centre (UK)
 - OpenDMP (<https://opendmp.eu/>) – OpenAIRE e EUDAT (UE)
 - RDMO (<https://rdmorganiser.github.io/en/>) – DFG (Alemanha)

Em resumo, propõem-se dois tipos de soluções para diferentes cenários de desenvolvimento do serviço nacional. Na primeira solução, de utilização de serviço na cloud, sugere-se a adesão a um modelo de serviço pré-estabelecido em que a entidade nacional deverá negociar com o proprietário da aplicação o pacote de serviço a usufruir e respetivos custos. As duas soluções são adequadas para o estabelecimento do serviço nacional, sendo que o sistema ARGOS, por via da integração da infraestrutura OpenAIRE, possui já integração da informação dos projetos da FCT. Por outro lado, o sistema DMPonline, apresenta pacotes de serviços e custos associados pré-definidos, com uma comunidade de utilizadores institucionais maior.

Na segunda solução, instalação local do software, existem três diferentes programas *Open Source* que satisfazem as necessidades técnicas e administrativas para estabelecimento de um serviço nacional que apoie a elaboração de PGDs. O software que está na base do serviço DMPonline é aquele que apresenta uma comunidade de utilizadores mais robusta, como é visível nos casos de serviço nacionais analisado no presente relatório. Por outro lado, o software OpenDMP, em uso na instância ARGOS, resulta de um trabalho conjunto das infraestruturas OpenAIRE e EUDAT para oferecer um serviço de PGDs para as infraestruturas e investigação. O software RDMO resulta do desenvolvimento nos últimos três anos de um serviço nacional na Alemanha com financiamento da DFG.

No presente relatório procede-se também a uma breve análise de serviços de âmbito nacional, com o objetivo de complementar e validar a conformidade das soluções propostas. O tipo de serviços e iniciativas nacionais consideradas na nossa análise (Espanha, Bélgica, França, Noruega, Dinamarca e Alemanha) confirma a pertinência de um serviço de âmbito nacional no apoio direto a políticas dos financiadores, sendo que quatro dos casos analisados adotaram o DMPonline e outros dois (Noruega e Alemanha) fizeram desenvolvimentos próprios de ferramentas.

Finalmente, infere-se do presente estudo que existe uma crescente consciencialização da importância do papel dos PGDs no quadro das políticas dos financiadores, mas igualmente na oferta de serviços de apoio à gestão de dados no desenvolvimento das infraestruturas de investigação. Tendo como referência as iniciativas nacionais identificadas e analisadas neste relatório, concluiu-se que existe a necessidade de oferecer um serviço de âmbito nacional e de desenvolver planos de formação no domínio dos PGDs, destacando-se o Fórum GDI como espaço privilegiado de dinamização da comunidade.